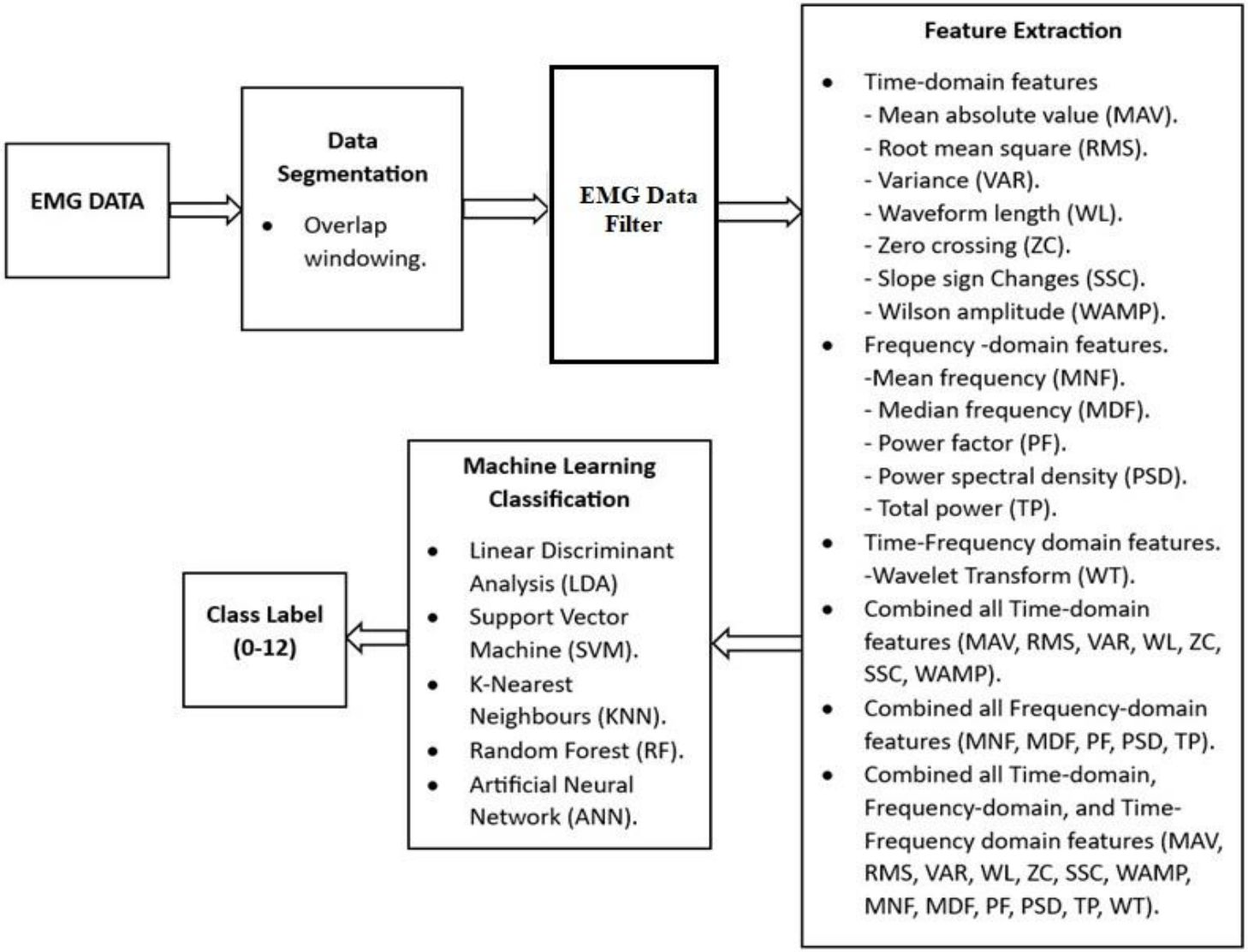
EEE8097: Key Findings

Contents

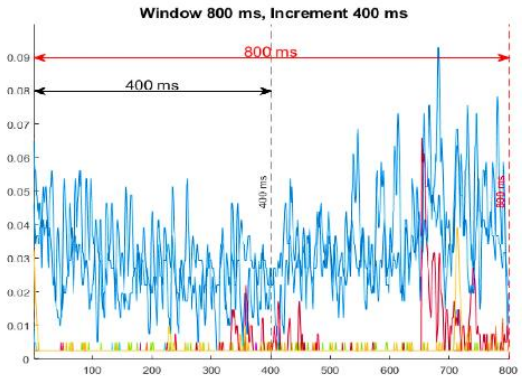
System Architecture

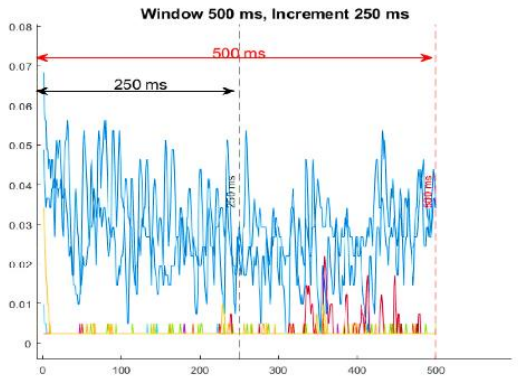


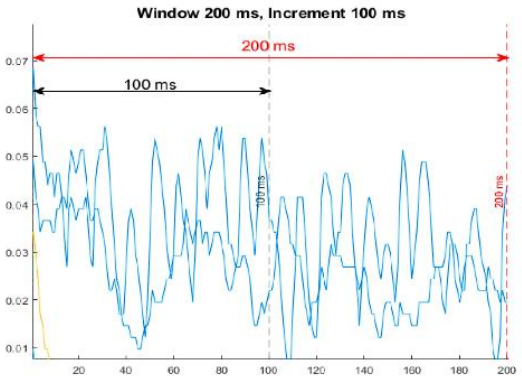
1. Experiment to find the best window size.

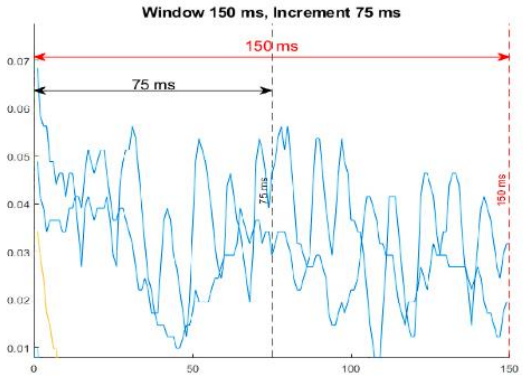
Window size

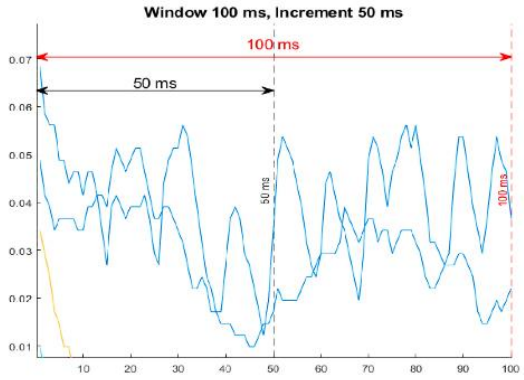
1. **800 - 400**
2. **500-250**
3. **200-100**
4. **150-75**
5. **100-50**
6. **50-25**

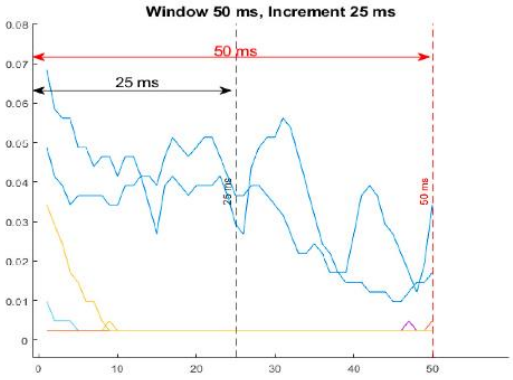












Note:

if winsize = 500 is meaned by each window will contain 500 samples from the signal.

Window increment size (sliding window) is specified how much window moves forward for the next segment.

For example,

if wininc = 250 is meaned by each sliding or overlapping window will start 250 samples after the previous window's start.

The number of windows (numwin) in each class (0 to 12) is calculated using below equation (1). Where N is the total number of samples in respective class. The '+1' ensure the initial window at the start of the signal, without including it miss the first segment window. Finally,

$$numwin = \left[\frac{N-winzise}{wininc}\right] + 1 \qquad (1)$$

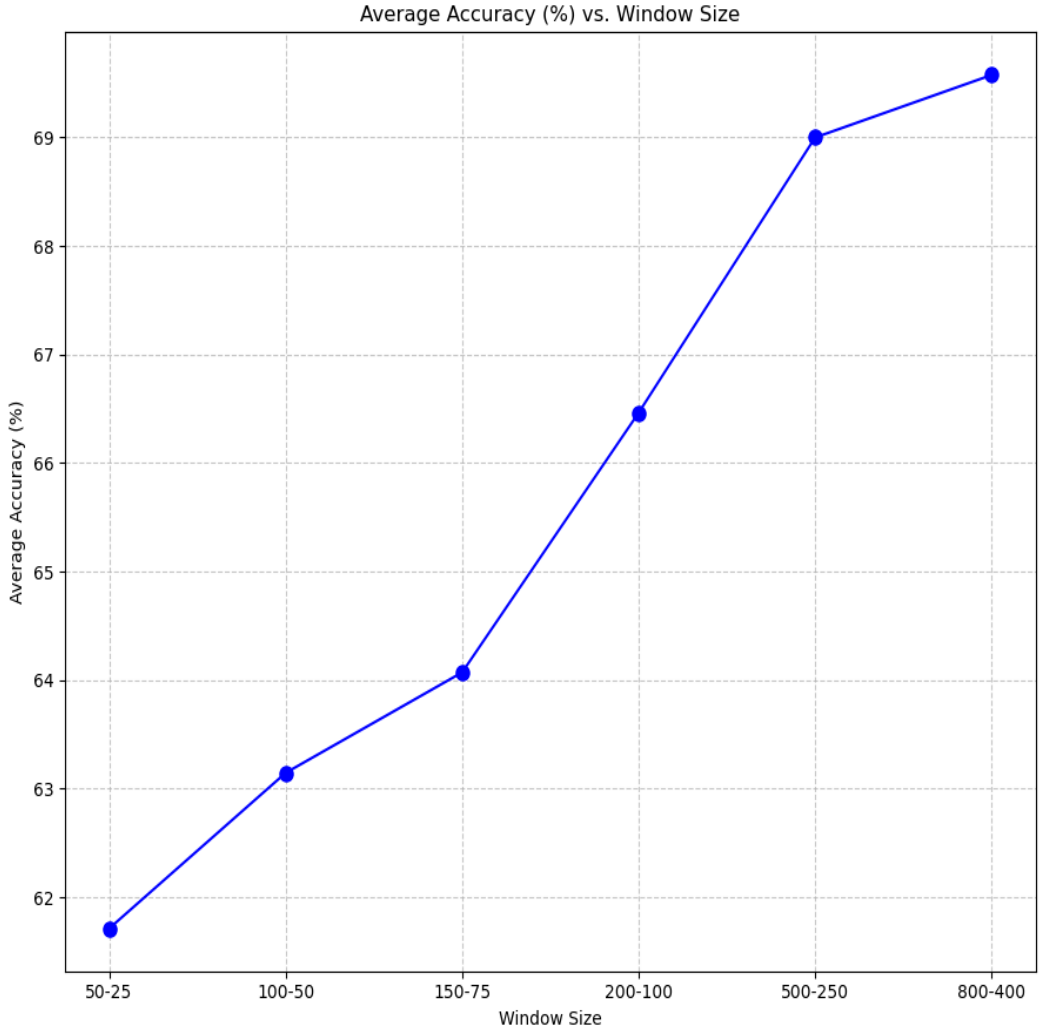
The window size in milliseconds depends on the sampling rate of the signal, here Ninapro DB1 experiment is used 10 Otto Bock MyoBock 13E200 electrodes [13].

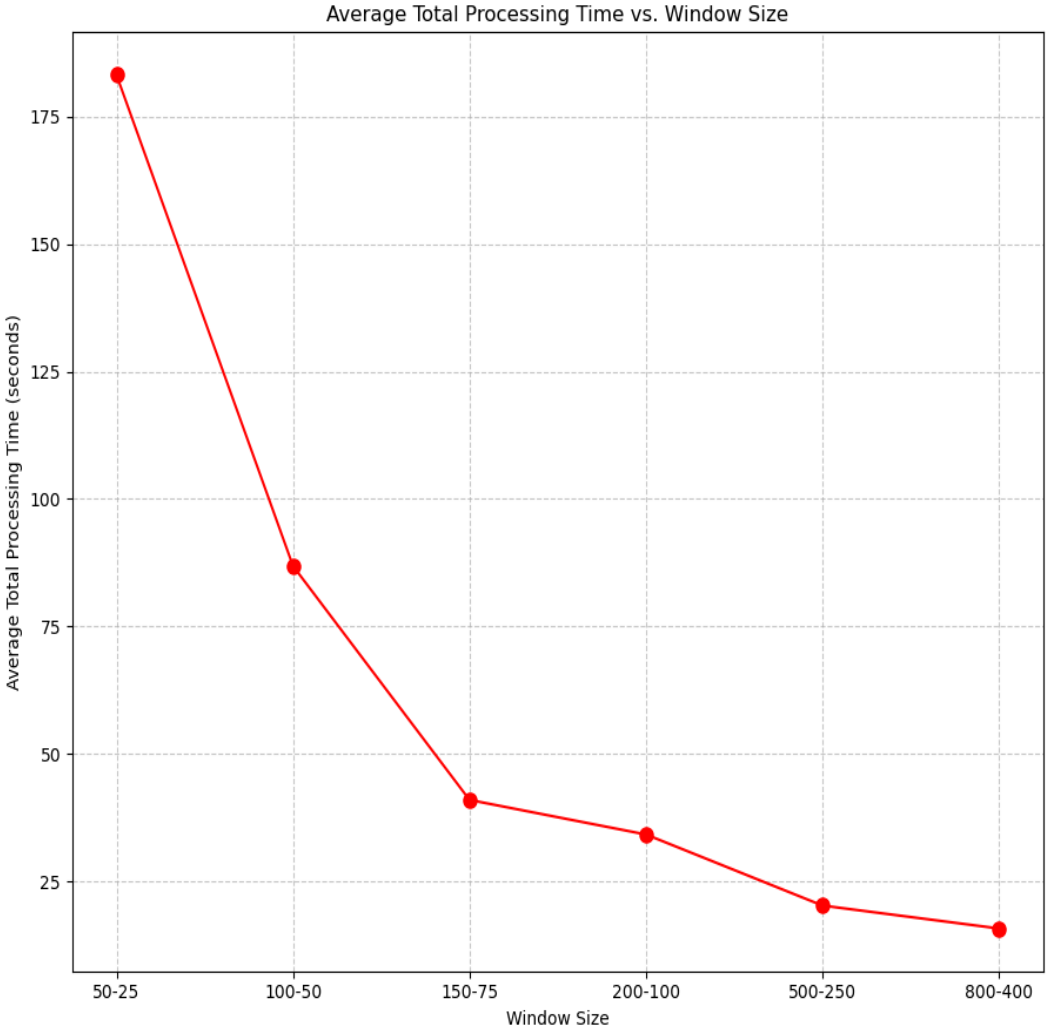
$$T = \frac{winsize}{Fs} \qquad (2)$$

a. Window Size v/s Accuracy and Processing Time.

Window size	Parameter	LDA	SVM	KNN	RF	ANN	Avg. Total
800-400	AVG. Accuracy	60.51	63.19	73.69	77.90	72.59	69.58
	AVG. Processing time	12.28	12.72	12.32	12.94	28.54	15.76
500-250	AVG. Accuracy	57.80	60.88	75.50	78.46	72.94	69.12
	AVG. Processing time	14.72	16.99	14.85	18.26	36.49	20.26
200-100	AVG. Accuracy	56.52	58.48	71.94	75.38	70.00	66.46
	AVG. Processing time	18.20	35.17	18.64	27.83	71.15	34.20
150-75	AVG. Accuracy	55.10	56.69	67.14	73.39	68.04	64.07
	AVG. Processing time	20.30	53.60	20.59	33.02	77.39	40.98
100-50	AVG. Accuracy	53.90	55.08	68.32	72.25	66.18	63.15
	AVG. Processing time	20.81	109.55	21.53	169.36	113.05	86.86
50-25	AVG. Accuracy	52.77	53.64	66.61	71.24	64.30	61.71
	AVG. Processing time	31.93	589.42	34.42	71.93	188.09	183.16

Plot.1 Window Size v/s Accuracy and processing time.





b. Key Findings.

- ✓ If window size increases, Accuracy of the model is increased.
- ✓ If window size increases. Processing time decreased.
- ✓ If window size is increased, Size of the dataset is decreased.
- ✓ Best window size as per the experiment are 8100-400 and 500-250.
- ✓ For further experiment window size 500-250 is selected due to below reasons
    1. Large window size (800-400) capable to capture long-term trends and overall patterns of the signals, but they might lead to underfitting by missing short-term significant events or temporal resolution of the signal.
    2. Smaller window sizes (50-25) are providing higher temporal resolution and capture the rapid changes in the signals, but they might lead to overfitting and higher noise content.
    3. Window size 500-250 is a trade of between accuracy and processing time of the model.

c. Contribution.

- ✓ Experimenting different window size and find the optimised window size is 500-250.

2. Experiment to find the best filter conditions.

a. General understanding Stimulus, Repetition, Restimulus and Rerepetition.

```
Combined data size of 1 to 27 Participiants: [2731393 12]
  EMG data_all size: [2731393 10]
  Stimulus_data_all size: [2731393 1]
  Repetition_data_all size: [2731393 1]
  Restimulus_data_all size: [2731393 1]
  Rerepetition_data_all size: [2731393 1]
Unique values in stimulus_data_all:
    0
    1
    2
    3
    4
    5
    6
    7
    8
    9
   10
   11
   12
```

```
Unique values in restimulus_data_all:
    0
    1
    2
    3
    4
    5
    6
    7
    8
    9
   10
   11
   12
```

```
For stimulus class 0:
  Unique values in repetition_data_all:
    0
  Number of unique values in repetition_data_all: 1
For stimulus class 1:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 2:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 3:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 4:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 5:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 6:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 7:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 8:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 9:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 10:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 11:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For stimulus class 12:
  Unique values in repetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
```



```
For restimulus class 0:
  Unique values in rerepetition_data_all:
    0
  Number of unique values in repetition_data_all: 1
For restimulus class 1:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 2:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 3:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 4:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 5:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 6:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 7:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 8:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 9:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 10:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 11:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
For restimulus class 12:
  Unique values in rerepetition_data_all:
    1    2    3    4    5    6    7    8    9    10
  Number of unique values in repetition_data_all: 10
```

b. Various Filter conditions v/s Mean Squared Error, Signal-to-Noise Ratio and Testing accuracy.

Filter Conditions	Average MSE (Original)	Average MSE (Filtered)	Average SNR (Original)	Average SNR (Filtered)	Classifiers	Testing Accuracy	Training Accuracy	Weighted Avg F1-Score
Model 1 (Stimulus filtered by Repetition condition 1& 2)	0.123535	0.200242	1.027378 dB	1.413006 dB	LDA	35.08%	36.54%	0.340
					KNN	85.25%	90.32%	0.850
					SVM	40.25%	40.89%	0.390
					RF	89.33%	100.00%	0.890
					Average	**62.48%**	**66.94%**	**0.618**
Model 2 (Stimulus filtered by Repetition condition 5,6,7,8,9,10)	0.123535	0.194971	1.027378 dB	1.361528 dB	LDA	53.44%	53.61%	0.480
					KNN	86.68%	92.13%	0.860
					SVM	56.13%	56.72%	0.510
					RF	91.11%	100.00%	0.910
					Average	**71.84%**	**75.62%**	**0.690**
Model 3 (Stimulus filtered by all Repetition condition (1to 10 and including 0))	0.123535	0.123535	1.027378 dB	1.027378 dB	LDA	53.44%	53.61%	0.480
					KNN	86.68%	92.13%	0.870
					SVM	56.13%	56.72%	0.510
					RF	91.11%	100.00%	0.910
					Average	**71.84%**	**75.62%**	**0.693**
Model 4 (Stimulus = Restimulus filtered by Repetition value = Rerepition)	0.123535	0.168031	1.027378 dB	1.267682 dB	LDA	60.70%	62.07%	0.560
					KNN	88.15%	92.25%	0.880
					SVM	62.96%	65.06%	0.590
					RF	89.43%	100.00%	0.890
					Average	**75.31%**	**79.85%**	**0.730**

Note:

Steps

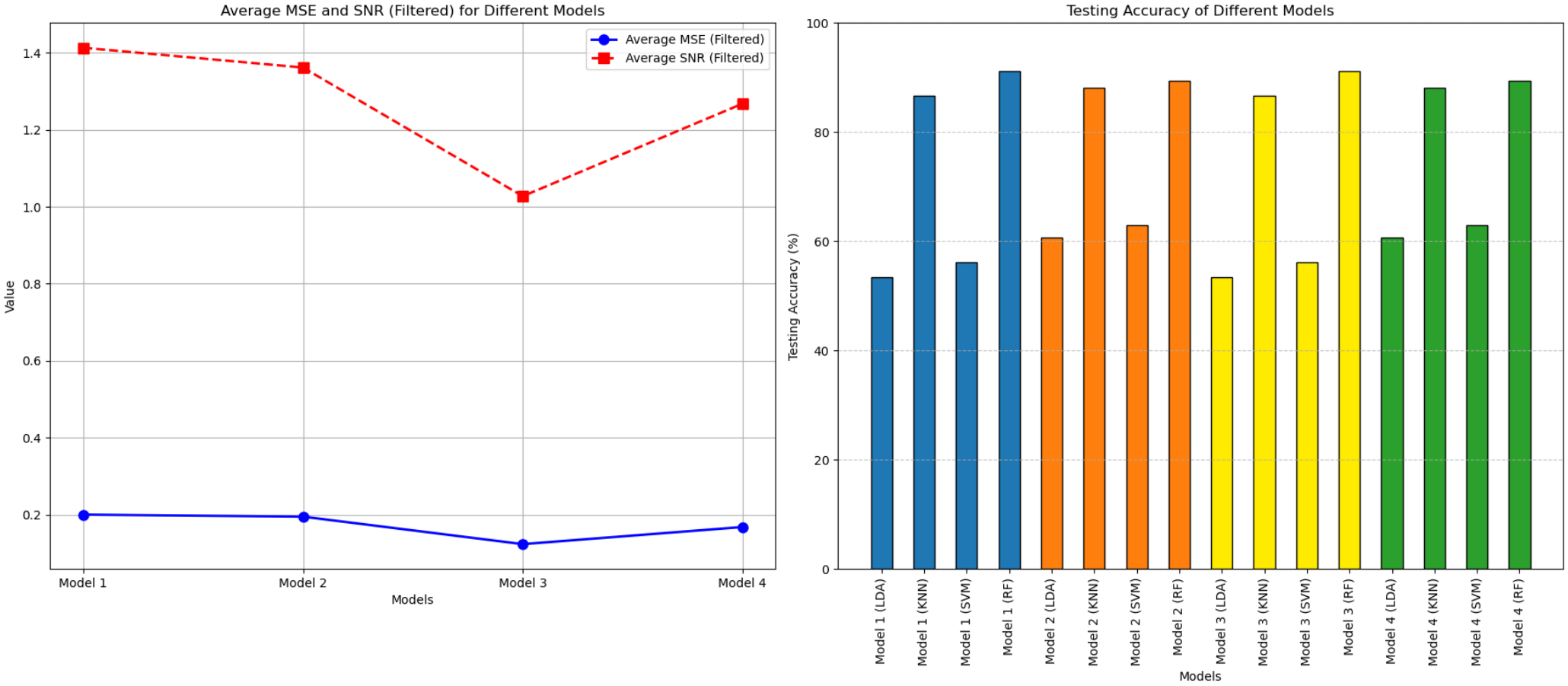
1. Find the MSE and SNR values for EMG signals before and after filtering.
   Definition: MSE is a measure of the average squared difference between the original (desired) signal and the filtered (or estimated) signal.
   Definition: SNR is a measure of the ratio of the power of the original signal to the power of the noise. It is often used to quantify the improvement in signal quality after filtering.

   MSE: Lower MSE indicates that the filtered signal is closer to the original signal, meaning the filter is more effective.

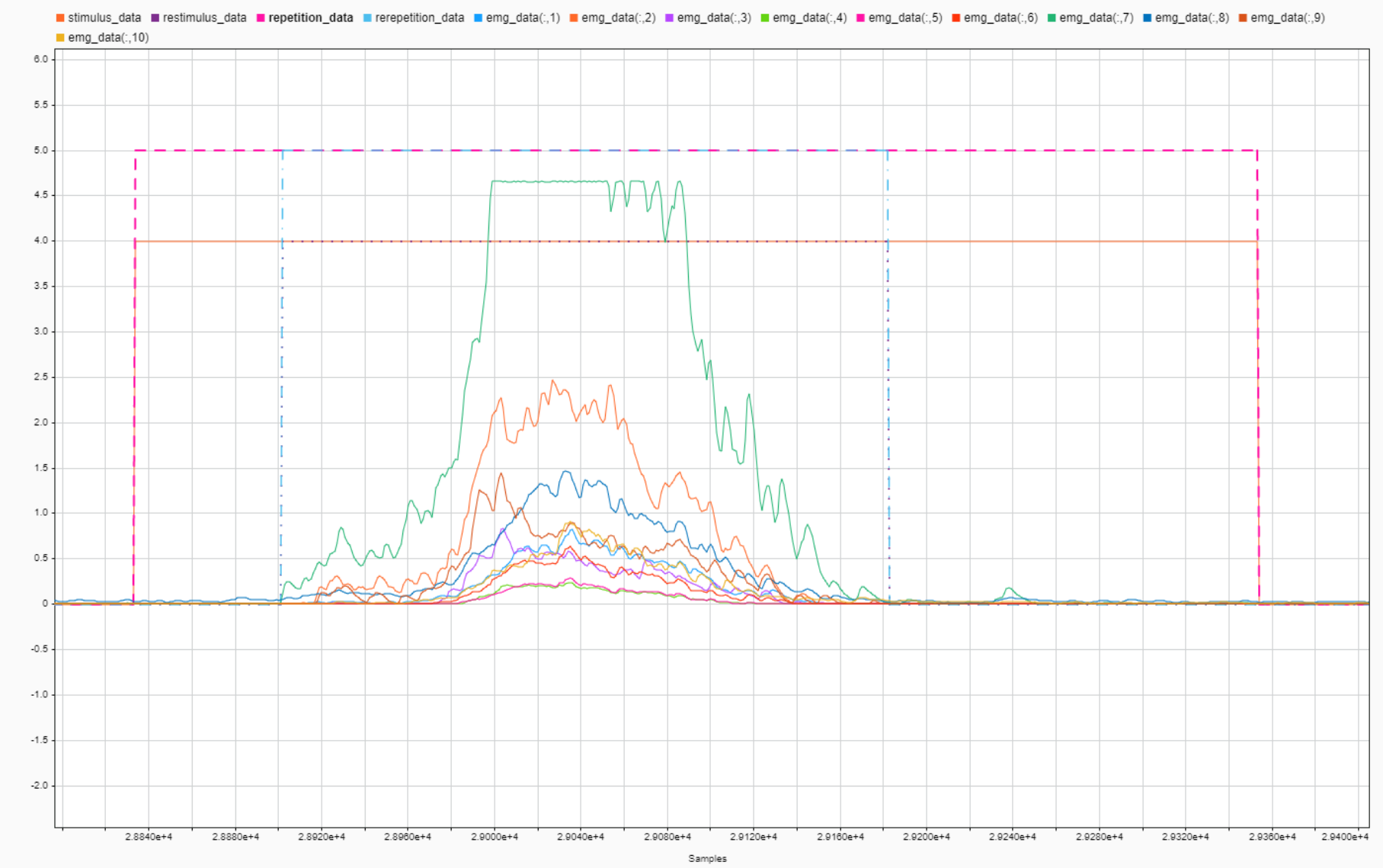
   SNR: higher SNR indicates a cleaner signal with less noise relative to the original signal.

2. Calculate the testing accuracy in each filter condition with various ML models (LDA, SVM, KNN & Random-Forest).

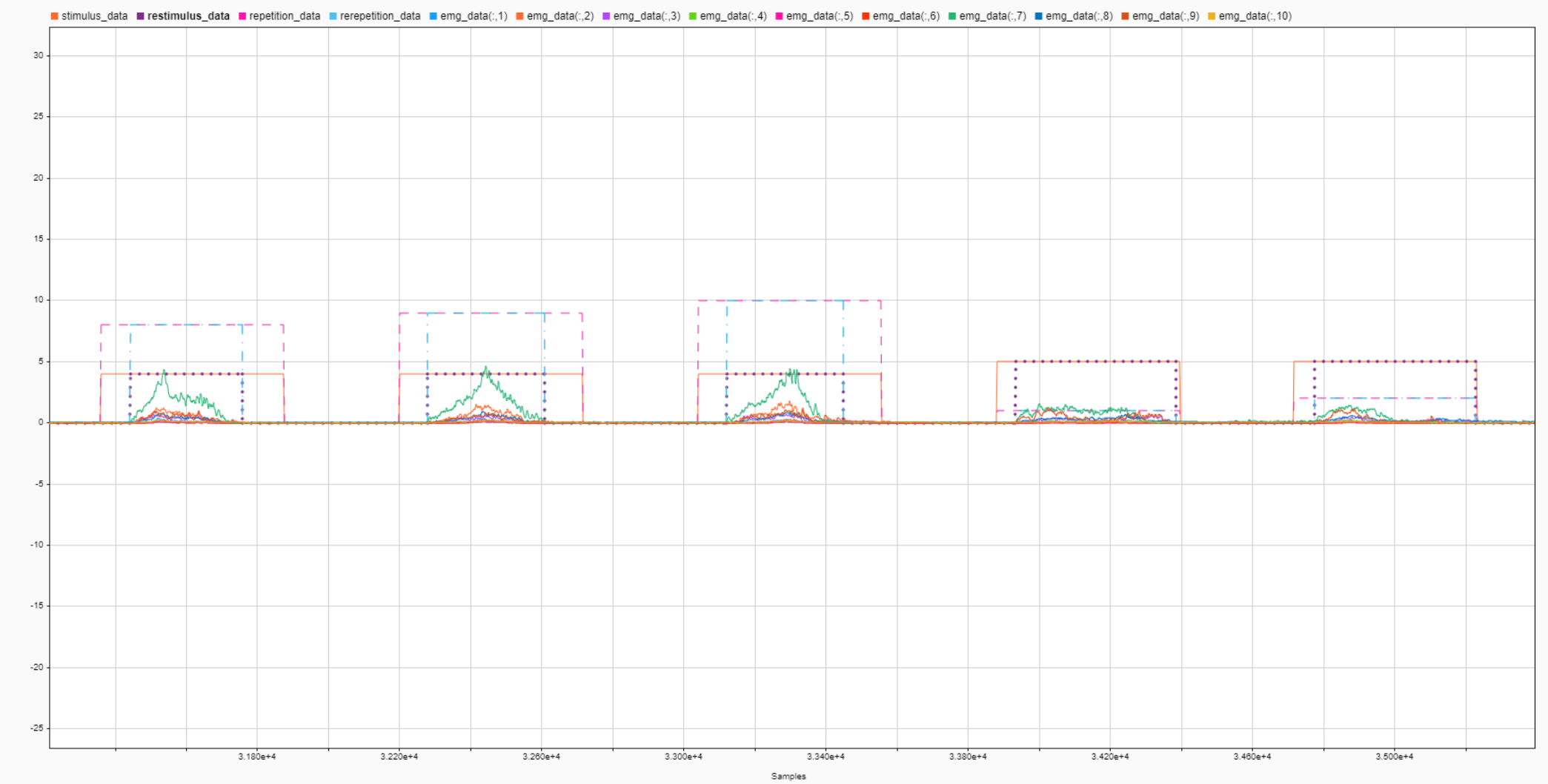
Plot. Graphical representation for average MSE and SNR with their accuracy.



Plot. Representation of 10-Channel EMG signals in Stimulus, Repetition, Restimulus and Rerepetition condition.



Plot. Multiple condition of Stimulus, Repetition, Restimulus and Rerepetition conditions.



d. Key findings

- ✓ Stimulus with increasing number of repetition, MSE and SNR are increased.
- ✓ Model:3 is just for testing condition for evaluate MSE and SNR in before and after filtering is true or not.
- ✓ Model 4 is the best filtering condition in terms of MS , SNR and Testing accuracy
- ✓ Best filter condition is a trade-off between MSE, SNR and Accuracy.

e. Contribution.

- ✓ Experiment different filter conditions and find the best condition for contribution low MSE, Higher SNR and Accuracy (Value of Stimulus, Restimulus, Repetition and Rerepetition are equal)

3. Combined Datasets

a. Find the best features using Statistical analysis (Mean, Median and Standard deviation).

STEPS

1. Calculate Mean, Median and Standard deviation for each channel with respective features.
2. Calculate average mean, average median, and average standard deviation in each feature.

Note: Mean is measured the central tendency of the feature values
Median shows middle value when data is stored, its helping to understand data distribution.
Standard deviation indicates the amount of variation or dispersion from the mean.

Mean: Measures the central tendency of the feature values.

Median: Indicates the middle value when the data is sorted, providing insight into the data distribution.

Standard Deviation: Reflects the amount of variation or dispersion from the mean.

Key Considerations for Feature Selection

1. **Feature Variability**: Features with appropriate variability (neither too low nor excessively high) can often contribute more useful information. Features with extremely low variability might not provide enough information for distinguishing between classes or predictions.
2. **Mean and Median Values**: While mean and median values can give insights into the central tendency of the features, what matters more for accuracy is how these features contribute to separating or predicting different classes or outcomes.
3. **Feature Distributions**: Features should have distributions that are suitable for the ML algorithm being used. For example, some algorithms perform better with normally distributed features.

Feature Comparison using Mean, Median and Standard deviation.

Dataset	Mean (Average of 10 channels)	Median (Average of 10 channels)	Standard Deviation (Average of 10 channels)
MAV Feature	0.17	0.08	0.24
RMS Feature	0.20	0.10	0.27
VAR Feature	0.04	0.00	0.13
WL Feature	4.71	2.93	5.35
ZC Feature	0.01	0.00	0.11
SSC Feature	153.77	157.10	76.19
WAMP Feature	4.38	0.20	13.91
MNF Feature	0.09	0.08	0.04
MDF Feature	0.05	0.03	0.05
PF Feature	0.0004	0.00	0.0010
PSD Feature	0.0017	0.0009	0.0021
TP Feature	0.83	0.46	1.07
TF_Energy_Feature	389.57	223.09	485.15

3. Best features based on Mean, Median and Standard deviation.

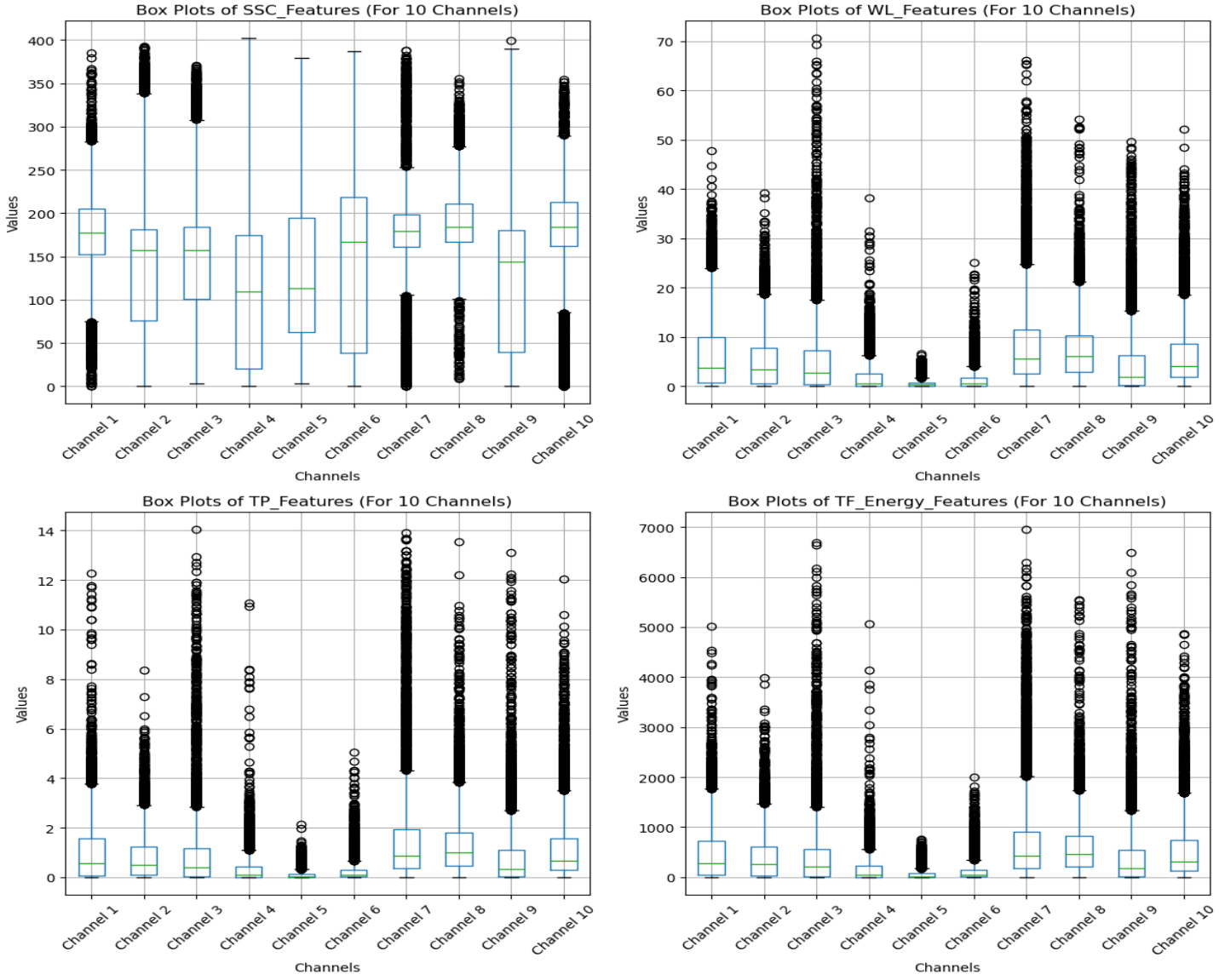
- **SSC Feature:** This feature stands out with its high mean, high median, and considerable standard deviation. It suggests that SSC captures a lot of variability and potentially holds significant discriminative power.
- **WL Feature:** high mean and median values while exhibiting significant variability,
- **TF_Energy_Feature:** The high standard deviation, large difference between mean and median, and potential presence of outliers or noise.

Category	Features
Best informative Features	SSC, WL, TP, TF_Energy
Moderate informative Features	WAMP, RMS, MNF, MDF, PSD
Less informative features	MAV, VAR, ZC, PF

b. Find the outliers in best features (SSC, WL, TP, and TF_Energy).

Channels	SSC Features	WL Features	TP Features	TF Energy Features
Outlier channel 1	676	295	335	318
Outlier channel 2	234	286	325	247
Outlier channel 3	252	357	461	445
Outlier channel 4	0	548	614	552
Outlier channel 5	0	594	730	640
Outlier channel 6	0	611	750	661
Outlier channel 7	1,144	545	683	683
Outlier channel 8	375	306	479	416
Outlier channel 9	1	616	638	587
Outlier channel 10	967	497	532	447

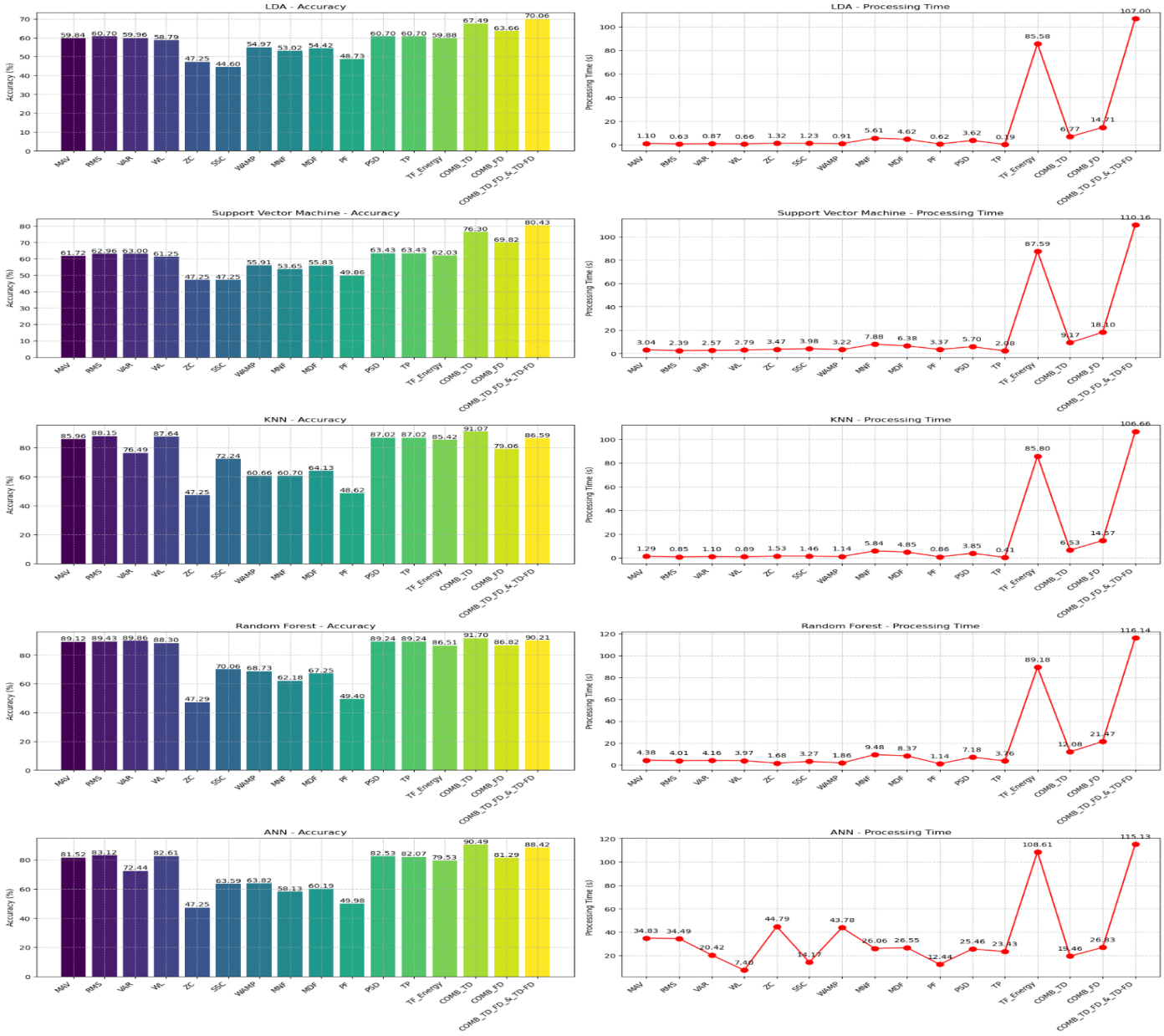
Box Plot representation of Beat features.



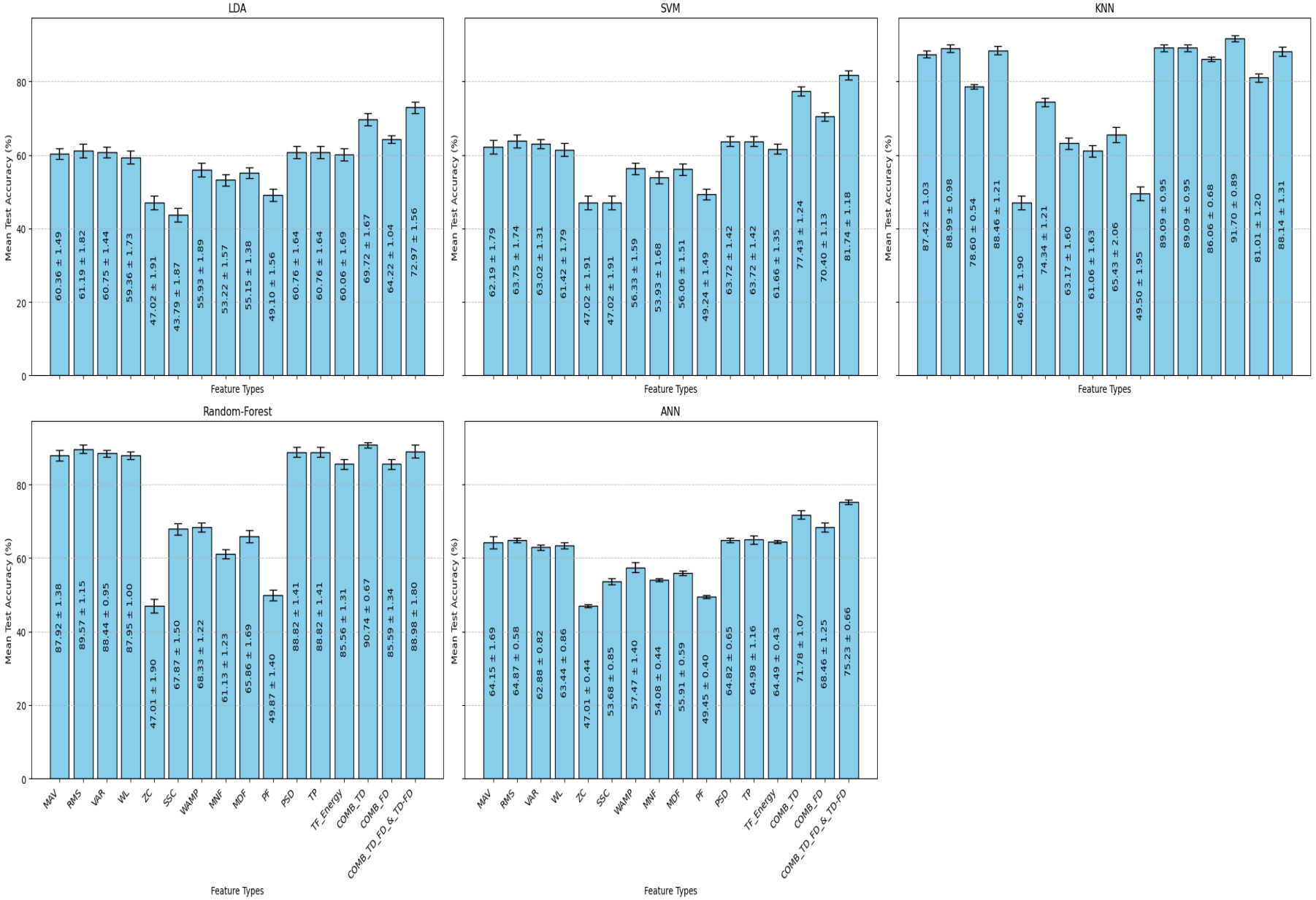
c. Find the best ML model.

1. Based on accuracy and processing time.

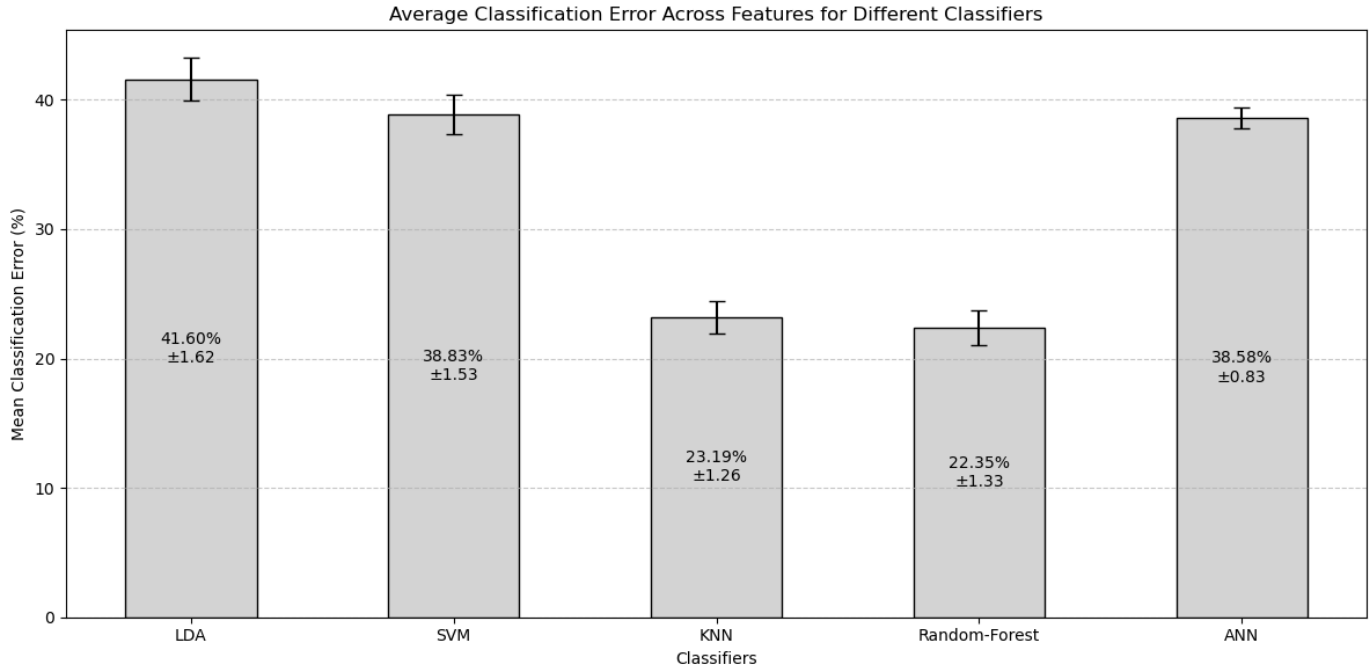
Window Size: 500-250										
Models	**LDA**		**SVM**		**KNN**		**Random Forest**		**ANN**	
Features	**Total Processing time(seconds)**	**Accuracy**	**Total Processing time(seconds)**	**Accuracy**	**Total Processing time(seconds)**	**Accuracy**	**Total Processing time(seconds)**	**Accuracy**	**Total Processing time(seconds)**	**Accuracy**
MAV	1.1	59.84	3.04	61.72	1.29	85.96	4.38	89.12	34.83	81.52
RMS	0.63	60.7	2.39	62.96	0.85	88.15	4.01	89.43	34.49	83.12
VAR	0.87	59.96	2.57	63	1.1	76.49	4.16	89.86	20.42	72.44
WL	0.66	58.79	2.79	61.25	0.89	87.64	3.97	88.3	7.4	82.61
ZC	1.32	47.25	3.47	47.25	1.53	47.25	1.68	47.29	44.79	47.25
SSC	1.23	44.6	3.98	47.25	1.46	72.24	3.27	70.06	14.17	63.59
WAMP	0.91	54.97	3.22	55.91	1.14	60.66	1.86	68.73	43.78	63.82
MNF	5.61	53.02	7.88	53.65	5.84	60.7	9.48	62.18	26.06	58.13
MDF	4.62	54.42	6.38	55.83	4.85	64.13	8.37	67.25	26.55	60.19
PF	0.62	48.73	3.37	49.86	0.86	48.62	1.14	49.4	12.44	49.98
PSD	3.62	60.7	5.7	63.43	3.85	87.02	7.18	89.24	25.46	82.53
TP	0.19	60.7	2.08	63.43	0.41	87.02	3.76	89.24	23.43	82.07
TF_Energy	85.58	59.88	87.59	62.03	85.8	85.42	89.18	86.51	108.61	79.53
Combined all Time-domain features (MAV, RMS, VAR, WL, ZC, SSC, WAMP)	6.77	67.49	9.17	76.3	6.53	91.07	12.08	91.7	19.46	90.49
Combined all Frequency-domain features (MNF, MDF, PF, PSD, TP)	14.71	63.66	18.1	69.82	14.57	79.06	21.47	86.82	26.83	81.29
Combined all Time, Frequency, Time-Frequency-domain features (MAV, RMS, VAR, WL, ZC, SSC, WAMP, MNF, MDF, PF, PSD, TP, WT)	107	70.06	110.16	80.43	106.66	86.59	116.14	90.21	115.13	88.42
Average	14.72	57.80	16.99	60.88	**14.85**	**75.50**	18.26	78.46	36.49	72.94


LDA - Accuracy
Accuracy (%)
59.84 60.70 59.96 58.79 47.25 44.60 54.97 53.02 54.42 48.73 60.70 60.70 59.88 67.49 63.66 70.06
LDA - Processing Time
Processing Time (s)
1.10 0.63 0.87 0.66 1.32 1.23 0.91 5.61 4.62 0.62 3.62 0.19 85.58 6.77 14.21 107.00
Support Vector Machine - Accuracy
Accuracy (%)
61.72 62.96 63.00 61.25 47.25 47.25 55.91 53.65 55.83 49.86 63.43 63.43 62.03 76.30 69.82 80.43
Support Vector Machine - Processing Time
Processing Time (s)
3.04 2.39 2.57 2.79 3.47 3.98 3.22 7.88 6.38 3.37 5.70 2.08 87.59 9.17 18.10 110.16
KNN - Accuracy
Accuracy (%)
85.96 88.15 76.49 87.64 47.25 72.24 60.66 60.70 64.13 48.62 87.02 87.02 85.42 91.07 79.06 86.59
KNN - Processing Time
Processing Time (s)
1.29 0.85 1.10 0.89 1.53 1.46 1.14 5.84 4.85 0.86 3.85 0.41 85.80 6.53 14.57 106.66
Random Forest - Accuracy
Accuracy (%)
89.12 89.43 89.86 88.30 47.29 70.06 68.73 62.18 67.25 49.40 89.24 89.24 86.51 91.70 86.82 90.21
Random Forest - Processing Time
Processing Time (s)
4.38 4.01 4.16 3.97 1.68 3.27 1.86 9.48 8.37 1.14 7.18 3.76 89.18 12.08 21.47 116.14
ANN - Accuracy
Accuracy (%)
81.52 83.12 72.44 82.61 47.25 63.59 63.82 58.13 60.19 49.98 82.53 82.07 79.53 90.49 81.29 88.42
ANN - Processing Time
Processing Time (s)
34.83 34.49 20.42 7.60 44.79 14.17 43.78 26.06 26.55 12.44 25.46 23.43 108.61 19.46 26.83 115.13
MAV RMS VAR WL ZC SSC WAMP MNF MDF PF PSD TP TF_Energy COMB_TD COMB_FD COMB_TD_FD_&_TD-FD

2. Based on Mean accuracy and standard deviation



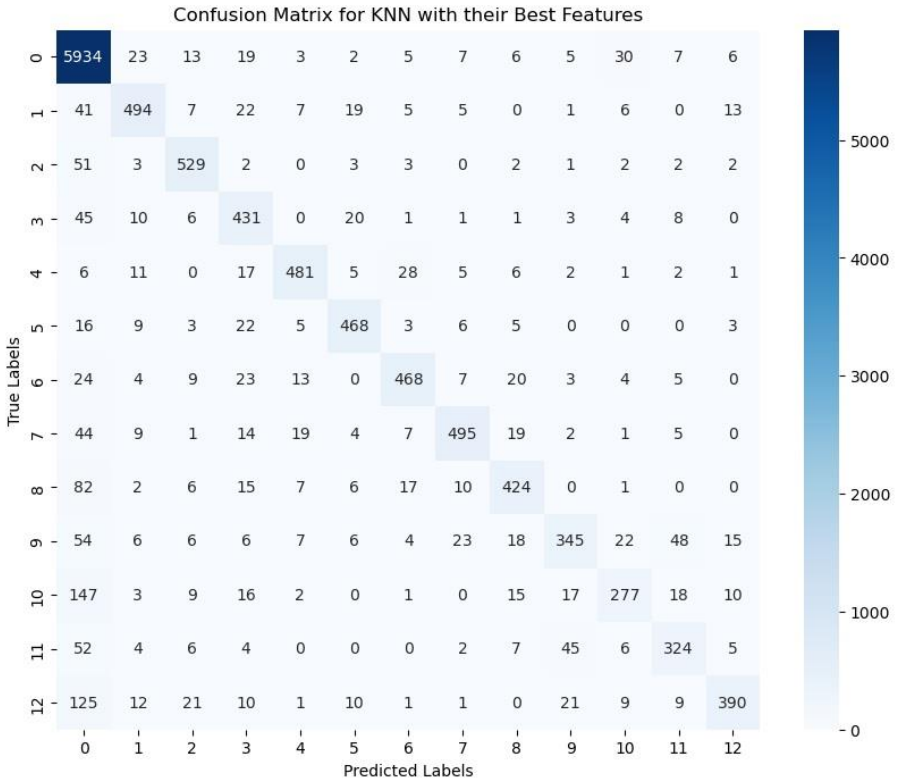
3. Based on classification error.

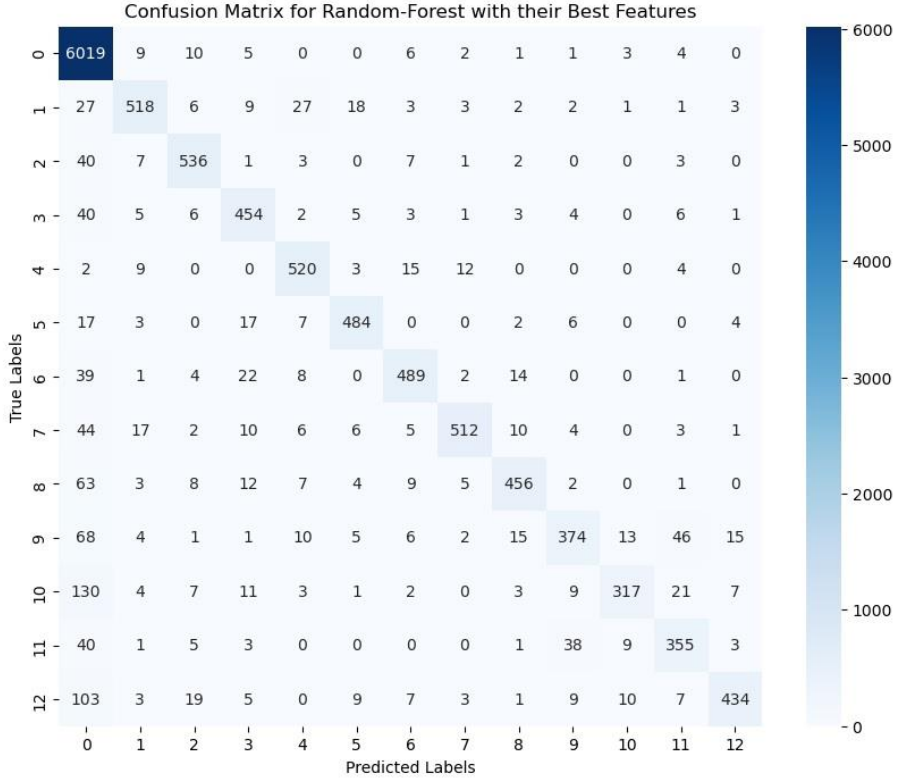


Results

Top Models and Their Best Features					
KNN	**Mean Accuracy (%)**	**Std Deviation**	**Random-Forest**	**Mean Accuracy (%)**	**Std Deviation**
Combined Time-Domain Feature	91.70	0.89	Combined Time-Domain Feature	90.74	0.67
Power Spectral Density Feature (PSD)	89.09	0.95	Root Mean Square Feature (RMS)	89.57	1.15
Total Power Feature (TP)	89.09	0.95	Combined Time, frequency, and Time-Frequency Domain	88.98	1.80
Wavelength Feature (WL)	88.82	1.41	Total Power Feature (TP)	88.82	1.41
Mean Absolute Value (MAV)	87.42	1.03	TF Energy Feature	88.82	1.31

4. Confusion matrix for best models and their features (KNN and Random-Forest)





f. Key Finding.

1. Best features (SSC, WL, TP, TF_Energy).
2. Outliers in best features (SSC, WL, TP, TF_Energy).
3. Best ML model KNNand Random-Forest based on Accuracy, Processing time, Mean Accuracy, Standard Deviation, Classification Error and Confusion Matrix.

g. Contributions

- ✓ Find the best Features and ML models by experimenting various methodes.

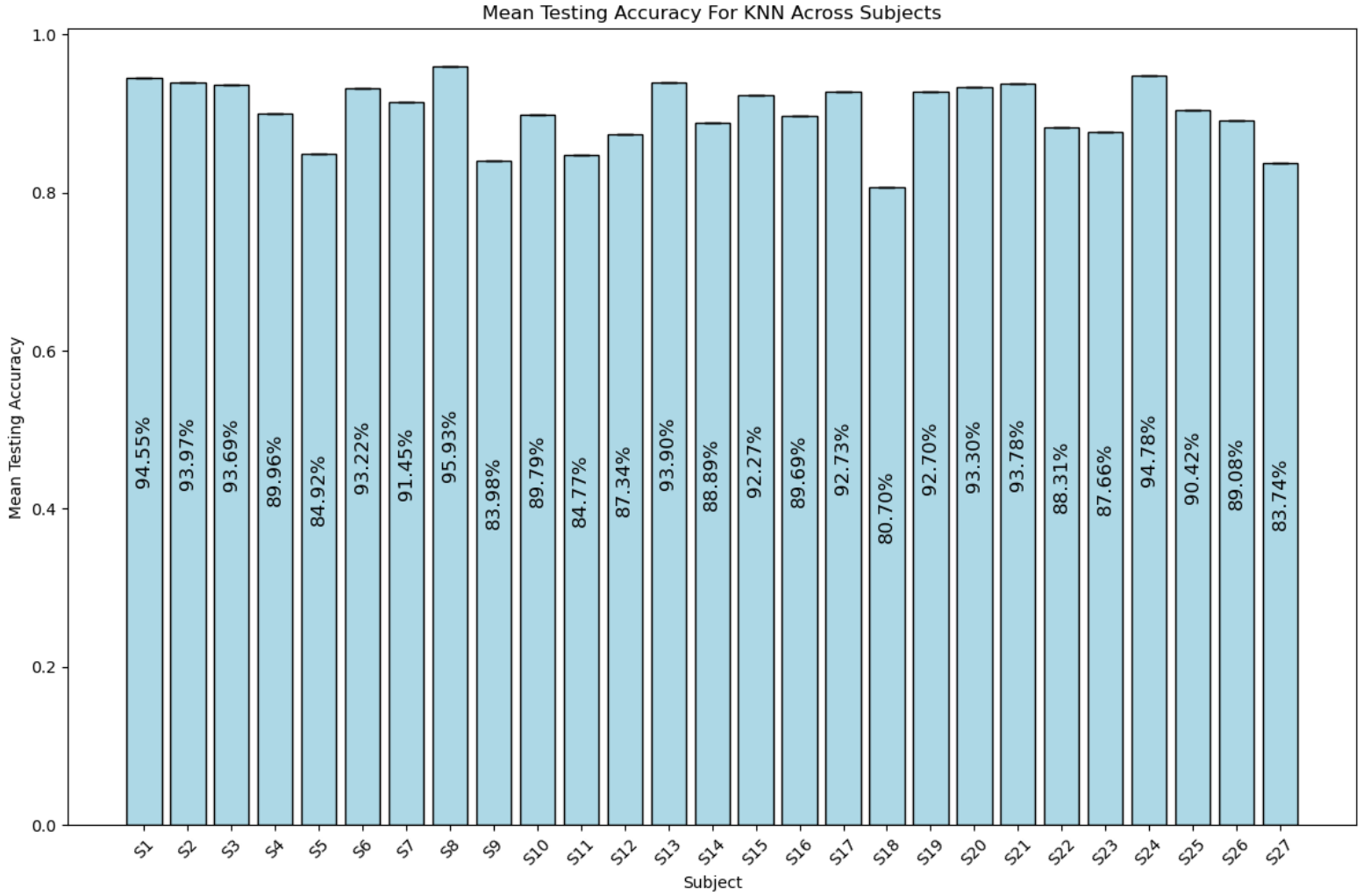
4. Individual dataset

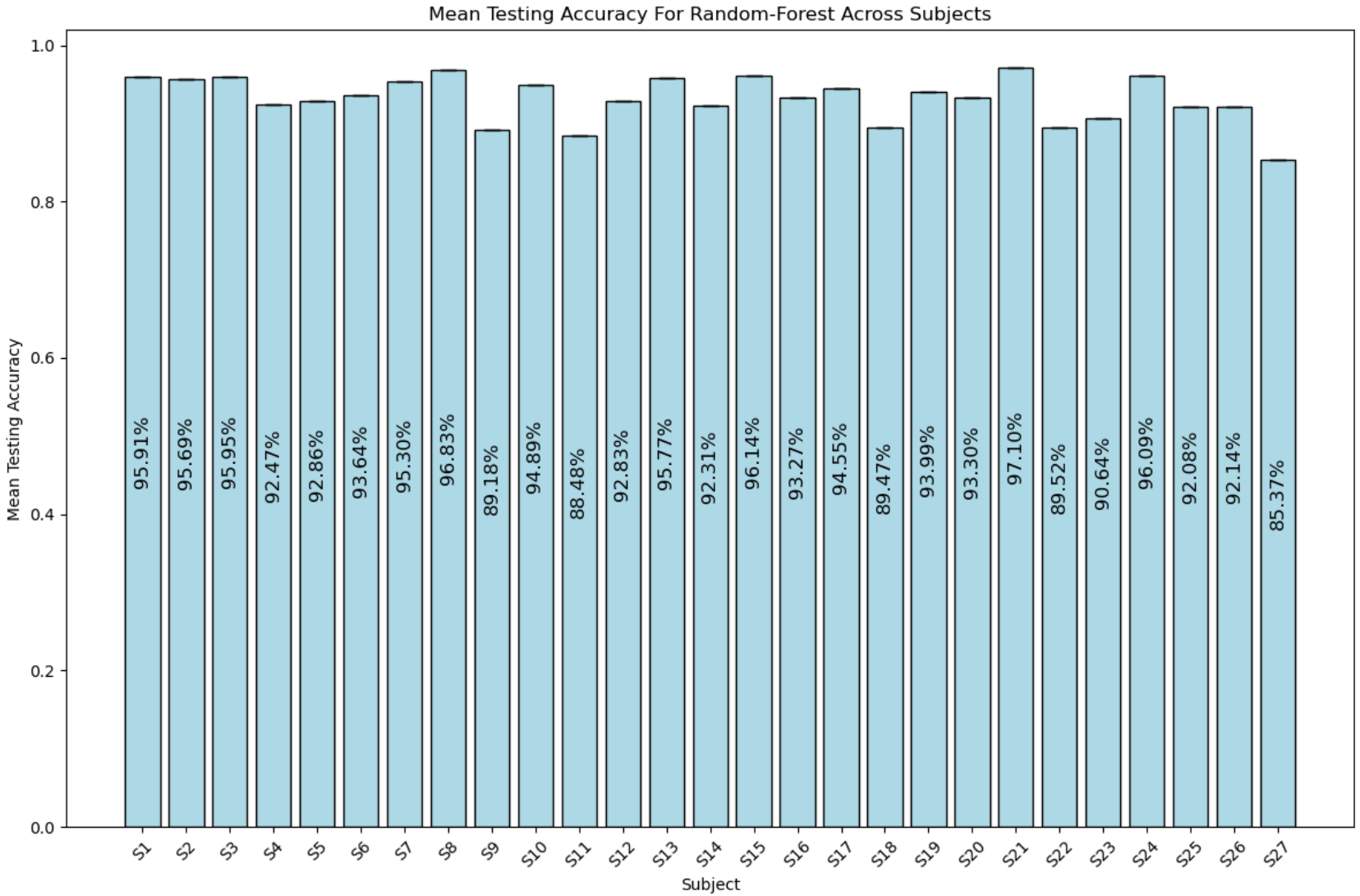
a. Find the best features-based on accuracy with KNN and Random-Forest.

Note: First calculate testing accuracy for all features and find the best features in all subjects.

Subjects	Features name	Random forest Testing Accuracy	KNN Testing Accuracy
S1	Combined_Time_and_Frequency_Domain_Feature	95.91	94.55
S2	Combined_Time_and_Frequency_Domain_Feature	95.69	93.97
S3	Combined_Time_and_Frequency_Domain_Feature	95.95	93.69
S4	Combined_Time_and_Frequency_Domain_Feature	92.47	89.96
S5	Combined_Time_and_Frequency_Domain_Feature	92.86	84.92
S6	Combined_Time_and_Frequency_Domain_Feature	93.64	93.22
S7	Combined_Time_and_Frequency_Domain_Feature	95.30	91.45
S8	Combined_Time_and_Frequency_Domain_Feature	96.83	95.93
S9	Combined_Time_and_Frequency_Domain_Feature	89.18	83.98
S10	Combined_Time_and_Frequency_Domain_Feature	94.89	89.79
S11	Combined_Time_and_Frequency_Domain_Feature	88.48	84.77
S12	Combined_Time_and_Frequency_Domain_Feature	92.83	87.34
S13	Combined_Time_and_Frequency_Domain_Feature	95.77	93.90
S14	Combined_Time_and_Frequency_Domain_Feature	92.31	88.89
S15	Combined_Time_and_Frequency_Domain_Feature	96.14	92.27
S16	Combined_Time_and_Frequency_Domain_Feature	93.27	89.69
S17	Combined_Time_and_Frequency_Domain_Feature	94.55	92.73
S18	Combined_Time_and_Frequency_Domain_Feature	89.47	80.70
S19	Combined_Time_and_Frequency_Domain_Feature	93.99	92.70
S20	Combined_Time_and_Frequency_Domain_Feature	93.30	93.30
S21	Combined_Time_and_Frequency_Domain_Feature	97.10	93.78
S22	Combined_Time_and_Frequency_Domain_Feature	89.52	88.31
S23	Combined_Time_and_Frequency_Domain_Feature	90.64	87.66
S24	Combined_Time_and_Frequency_Domain_Feature	96.09	94.78
S25	Combined_Time_and_Frequency_Domain_Feature	92.08	90.42
S26	Combined_Time_and_Frequency_Domain_Feature	92.14	89.08
S27	Combined_Time_and_Frequency_Domain_Feature	85.37	83.74

b. Find the mean accuracy for combined time and frequency domain feature in each subject (1 to 27) with KNN and Random-Forest.

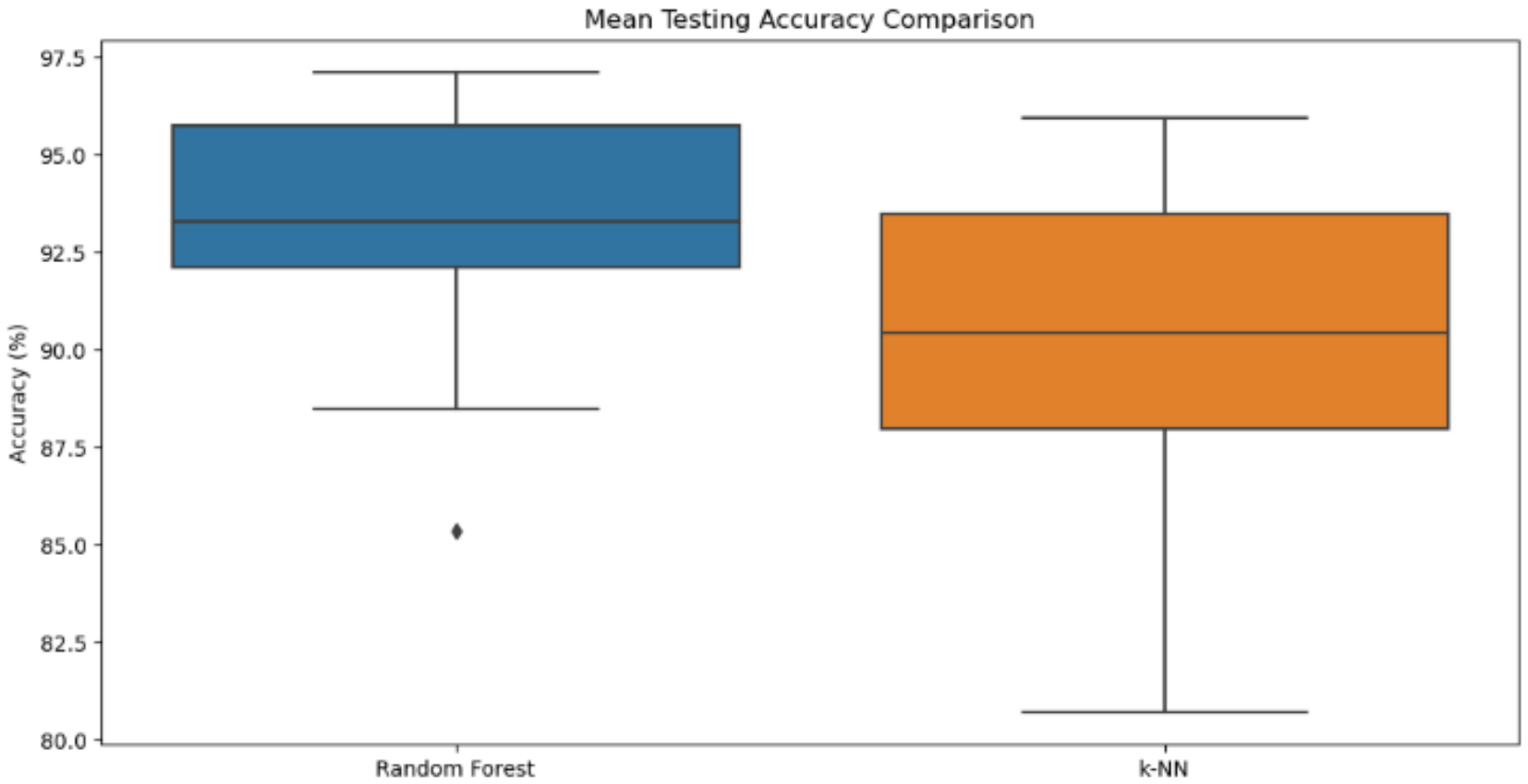



Mean Testing Accuracy For Random-Forest Across Subjects
Mean Testing Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
95.91%
95.69%
95.95%
92.47%
92.86%
93.64%
95.30%
96.83%
89.18%
94.89%
88.48%
92.83%
95.77%
92.31%
96.14%
93.27%
94.55%
89.47%
93.99%
93.30%
97.10%
89.52%
90.64%
96.09%
92.08%
92.14%
85.37%
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
S13
S14
S15
S16
S17
S18
S19
S20
S21
S22
S23
S24
S25
S26
S27
Subject

b. Find the best healthy subjects across 27.

✓ The top subjects based on accuracy in both classifiers are 1, 2, 3, 8, 13, 21 and 24.

c. Calculate mean testing accuracy for Random-forest and KNN.



d. Statistical test for performance evaluation of three different features (WL, TP, and TF-Energy) across 27 subjects.

	Testing accuracy			Ranking		
Subjects	**wl_feature**	**tp_feature**	**tf_energy feature**	**wl_feature**	**tp_feature**	**tf_energy feature**
S1	95.45	93.18	S2	3	2	1
S2	92.24	95.26	S3	2	3	1
S3	95.95	95.95	93.69	2.5	2.5	1
S4	90.80	93.31	90.38	2	3	1
S5	88.10	93.25	84.52	2	3	1
S6	89.41	94.92	87.71	2	3	1
S7	93.16	94.44	90.17	2	3	1
S8	95.93	96.38	94.57	2	3	1
S9	87.88	88.31	83.55	2	3	1
S10	94.89	94.47	92.77	3	2	1
S11	87.24	89.71	83.95	2	3	1
S12	85.23	90.72	83.54	2	3	1
S13	92.49	96.71	92.49	1.5	3	1.5
S14	85.47	90.60	81.62	2	3	1
S15	93.56	95.71	90.99	2	3	1
S16	91.48	91.93	91.48	1.5	3	1.5
S17	95.00	95.45	95.00	1.5	3	1.5
S18	86.84	88.60	79.82	2	3	1
S19	90.13	93.56	87.98	2	3	1
S20	93.75	91.52	90.18	3	2	1
S21	95.44	95.02	95.44	2.5	1	2.5
S22	87.90	89.92	85.08	2	3	1
S23	86.81	91.06	86.38	2	3	1
S24	93.04	95.22	90.43	2	3	1
S25	91.25	93.75	90.00	2	3	1
S26	89.96	88.65	90.83	2	1	3
S27	81.71	87.40	83.33	1	3	2
Mean Ranking				**2.06**	**2.72**	**1.22**

e. key Findings.

1. The Combined time and Frequency domain feature is the best across 27 subjects.
2. Find the best healthy subjects across 27 subjects.
3. TF-Energy feature is the significant feature by statistical testing across three features (WL, TP, and TF-Energy)

f. Contributions

✓ Analysing various experiment and evaluation to find the best features and ML models in Individual subjects (Combined time and frequency domain, Random-Forest).

5. Hybrid Model.

a. Find the accuracy in two models (KNN and Random-Forest).

Steps

1.Choose best features in all domains (wavelength, total power, and time-frequency energy) and combine to single dataset and find the accuracy in different models (KNN and Random-Forest)

Metric	KNN	Random-Forest
Testing Accuracy	88.89%	90.49%
Training Accuracy	93.33%	100.00%
Macro Average Precision	0.85	0.9
Macro Average Recall	0.82	0.84
Macro Average F1-Score	0.83	0.86
Weighted Average Precision	0.89	0.91
Weighted Average Recall	0.89	0.9
Weighted Average F1-Score	0.89	0.9

b. Key Finding.

1. Data Size is reduced and increase the accuracy and processing time.
2. For getting higher accuracy and least computation time, no need to combine all feature jus need to combine the relevant features in different domains.

c. Contributions

- ✓ Developed optimised model based on higher accuracy and least computation time.